Summer 1980

Effects of Assessor Training on Subsequent Assessment Center Performance

Michael R. Struth
University of Central Florida

EFFECTS OF ASSESSOR TRAINING ON SUBSEQUENT ASSESSMENT CENTER PERFORMANCE

BY

MICHAEL R. STRUTH
B.A., University of West Florida, 1975

THESIS

Submitted in partial fulfillment of the requirements for the degree of Master of Science: Industrial Psychology in the Graduate Studies Program of the College of Social Sciences at University of Central Florida; Orlando, Florida

Summer Quarter
1980

ACKNOWLEDGEMENTS

The author would like to thank Drs. Wayne A. Burroughs, Fredric D. Frank, and Cabot L. Jaffee for their counsel and assistance in the planning and writing of this paper.

A very special thanks is extended to Fred Frank not only for assistance in the preparation of this paper, but also for his personal and professional counsel over the years. My association with him has been a continuing source of personal and professional growth which I greatly value.

TABLE OF CONTENTS

INTRODUCTION

The purpose of this study was to investigate the effects of in-depth experience with and exposure to the assessment center process, i.e., assessor training, on subsequent assessment center performance. A study of this issue was necessary for several reasons which are discussed below.

In recent years the growth of assessment center technology utilization by both industry and government has been substantial. Several years ago estimates were that approximately one hundred organizations were operating assessment centers (Huck, 1973). Today, estimates as to the number of users range upward of five hundred organizations (Cohen, 1978). Although Bender's (1973) survey of organizational users indicated that only a few companies utilized assessment centers at more than one organizational level, it seems safe to conclude that this statistic has also increased dramatically.

As the number of assessment center users has grown, so has the number of individuals who have been exposed to the assessment center process. This exposure is of two basic types: in one case an individual may experience the process in the role of an "assessee," i.e., the individual

evaluated by the process; in a second case an individual may experience the process in the role of an "assessor," i.e., the individual who serves as an evaluator of assessees. Thus, as the use of assessment centers has grown, the number of individuals exposed to the process through both of these roles has grown.

For organizations which utilize the assessment center process at more than one organizational level, some potentially serious issues exist. In such organizations it is quite possible that individuals who have been trained to serve as assessors for centers conducted at lower organizational levels may, at later times, find themselves in the role of an assessee for centers targeted at higher organizational levels. In such situations, organizations face two concerns. On one hand, there is a substantial chance that fellow assessees, who have not been previously trained as assessors, will voice concern regarding an unfair testing situation; concern that those assessees who have experienced assessor training have an unfair advantage because of such training. Such criticism of the process represents a political concern to the organization; however, at the heart of this issue are the effects of practice and training effects on assessment center performance.

If, in fact, individuals who have been previously trained as assessors score higher in assessment center

evaluations than do assessees who have not experienced such training, alternative explanations are possible. A possible interpretation could be that assessor training leads to improvement in the skills measured by the process. In other words, training having to do with the observation and evaluation of others' management skills could lead to improvement in the management skill abilities of the assessor. An equally plausible explanation could be that although assessor training can lead to improved assessment center performance, such improvement is the result of familiarity with the process itself rather than an increase in the skill abilities of such individuals. In other words, the assessor training experience could lead to improved performance in the assessment center process simply because such individuals are intimately familiar with behaviors which are evaluated positively by assessors, and can thus "fake" effective performance. To date there has been no research to determine what, if any, effect assessor training has on subsequent assessment center performance. Thus, the purpose of this study was to determine if, in fact, assessment center performance differences existed which could be attributed to assessor training.

Although no research has been published on this issue, some claims have been made regarding the benefits of assessor training. Byham (1970) indicated that one of the most

important "fringe" benefits derived by managers who are trained as assessors is improvement in their own management skills. While this appears to be a plausible claim, support for such a claim has not been demonstrated through on-the-job measures of managerial effectiveness nor through an assessment center measurement of managerial effectiveness. Of course there is a possibility that such individuals may improve in skills not measured by the assessment center process, such as improvements in interviewing skills or broadening of observation skills; however, the study presented here addressed only management skills measured by the assessment center process.

Research on the assessment center for the most part has been concerned with determining the reliability of the assessment center process in general (Bray, Campbell, and Grant, 1974) or the reliability of performance on individual simulation exercises (Moses, 1973). These studies did not attempt to investigate variance in performance across measurements as a function of prior assessment center experience; however, some investigators have addressed this issue.

Bass' research (1954) was one of the earliest investigations of the effects of coaching on performance in a simulation instrument similar in nature to those instruments used in present day assessment centers. To

investigate this issue, two groups of assessees were coached; one group having demonstrated relatively low performance in an initial leaderless group discussion exercise, the other group having demonstrated relatively high performance in an initial leaderless group discussion exercise. When both groups were then evaluated in a second group discussion, statistically significant improvement in performance emerged which was attributable to coaching effects; however, the effect differed for the two groups. Although significant performance improvement was demonstrated for the group which had initially performed relatively well in the pre-test group discussion measure, no significant improvement was demonstrated for the group which had performed relatively weaker in the pre-test group discussion. Though the results demonstrated statistically significant improvement for one of the coached groups, in an absolute sense the differences between groups were small and Bass concluded that the differences were too small to be of practical concern.

The subject of coaching as an influencer of performance on simulation exercises remained dormant after the initial studies of Bass until Petty (1974) attempted to expand on previous studies. Petty, also utilizing a leaderless group discussion exercise, assigned subjects to one of four conditions: no experience and no training; no

experience but training; experience but no training; experience and training. In this study the independent variable of experience was defined as participation or lack of participation in a leaderless group discussion exercise prior to participation in the leaderless group discussion through which dependent variable measurements were obtained. The independent variable of training consisted of subjects being or not being informed of the exact behavioral activities which would result in favorable performance evaluations on the dependent variable measures. Subjects were given a brief lecture on the history and relevant research on leaderless group discussion exercises and were also allowed to inspect the behavioral checklist which would be used to evaluate their performance in the exercise. The dependent measurements consisted of observer evaluations as to the degree to which a subject initiated structure in the group discussion, the amount of consideration a subject displayed toward other group members, and an overall evaluation of effectiveness of the subject in the leaderless group discussion. The results of the study demonstrated that significant performance improvements were attributable to both independent variables. Significant main effects were found for the variable of training on all three criteria measures; the variable of experience produced significant main effects on only the measure of overall

effectiveness in the group discussion. Although these findings were statistically significant, they were modest in terms of practical significance. Less than ten percent of the variance in the dependent measure was accounted for by the independent variables.

A similar study which addressed the effects of prior exposure to simulation instruments on subsequent performance on simulation instruments was the work of Burroughs, Rollins and Hopkins (1973). In this study experimental subjects were asked to observe and evaluate participants, in a video-taped leaderless group discussion exercise, according to a number of behavioral skill categories. Control subjects did not review the video-taped simulation exercise. Subsequently a mix of control and experimental subjects actually participated in a second leaderless group discussion exercise. During the second group discussion assessors, unaware of which participants were control versus experimental subjects, observed the group discussions and assigned overall performance ratings to each participant on the same skill categories with which experimental subjects had been familiarized. In addition, all group discussion participants were rank ordered according to their overall performance in the group discussion. Analysis of both the ratings and rankings of subjects failed to

demonstrate any significant performance differences between experimental and control subjects.

Another study similar in nature was the work of Jaffee and Michaels (1978) on the effect of coaching on in-basket performance. In this study, three groups of subjects were utilized: the control group which consisted of subjects who took the in-basket with no prior coaching on the exercise; an experimental group which consisted of subjects that received coaching from the control group immediately after the control group had taken the in-basket; an experimental group which consisted of subjects that received coaching from the control group one week after the control group had taken the in-basket. The results of the study demonstrated no significant differences between the experimental subjects and control subjects on the management skills assessed in the in-basket. As a result the investigators concluded that informal information which might be passed from one assessee to another during an actual assessment center cycle would not significantly affect in-basket performance.

Another study which attempted to investigate the issue of coaching effects was the work of Denning and Grant (1979). In this study, the effect of coaching on performance in a leaderless group discussion exercise was examined. Observations were obtained on twelve leaderless

group discussions with each group being comprised of three experimental subjects and three control subjects. Experimental subjects were given written material which contained information regarding the definitions of the skill categories on which they were to be evaluated, as well as, specific behaviors which would correspond positively to each skill category and a description of the scale on which they would be evaluated. Control subjects were provided with only a general description of the assessment center process at large; no specific information was supplied the control group regarding the leaderless group discussion exercise in which they would participate. The results of the study demonstrated no statistically significant differences between groups and the experimenters concluded that prior knowledge of the process nor coaching significantly influenced performance in the exercise.

The above referenced studies were similar in several respects: each dealt with coaching effects on subsequent performance in simulation instruments typically used in operational assessment centers; each utilized only a single simulation instrument in investigating these effects; each utilized brief coaching sessions as an independent variable; also, each failed to demonstrate considerable practical significance for the effects of coaching on subsequent performance on simulation instruments.

The study reported here represented a substantial departure from these studies. In this study, in-depth exposure and experience with the assessment center process was utilized, i.e., assessor training. Secondly, this study attempted to address the issue in a manner which was more generalizable, in the practical sense, to organizations now operating assessment centers by utilizing performance across a number of instruments, as opposed to performance on a single simulation instrument, as the dependent measure.

The study presented here utilized assessment center performance results from an actual, rather than laboratory, assessment center program; this being the 1977 assessment center performance of candidates for a government organization's executive development program. The study afforded a unique opportunity to investigate directly the effects of assessor training on subsequent assessment center performance since some candidates for the executive development program had been trained as assessors for assessment centers conducted at lower organizational levels while other candidates had received no such training. A summary matrix of simulation instruments and the management skills measured therein for the executive development program is contained in Appendix A.

Hypotheses

In line with the research of Burroughs, Rollins, and Hopkins (1973), Denning and Grant (1979), and Jaffee and Michaels (1978) no differences in the assessment center performance between the control and experimental groups resulting from the independent variable of assessor training was predicted. The specific hypotheses of the study are listed below.

1. No performance differences were predicted relative to the control and experimental groups' overall assessment center performance evaluations.
2. No differences were predicted relative to the control and experimental groups' performance on any of the overall skill category evaluations (i.e., leadership, sensitivity, decisiveness, organizing and planning, adaptability, perception, decision making, oral communication, and written communication).
3. No differences were predicted relative to the control and experimental groups' overall performance on any of the individual simulation instruments (i.e., in-basket, problem solving, competitive leaderless group discussion, cooperative leaderless group discussion, and leadership simulation instrument).

4. No interaction effects were predicted between the factors of treatment group and type of simulation instrument for any of the skill categories assessed.

Method

Subjects

All subjects in the study were applicants to an executive development program conducted by a government organization. The experimental group was comprised of applicants who had been previously trained and served as assessors for assessment centers conducted at lower organizational levels. The control group was comprised of applicants who had received no such training. Because the design of the study reported here was a post-test only, there was concern regarding the equivalency of the two groups with respect to their management abilities.

Selection of individuals to be trained and serve as assessor for other assessment programs was not interpreted as indicating that superior management skills were possessed by selected individuals. Rather, these selection decisions were a function of availability, receptiveness toward the assessment center process, willingness to work long hours during assessment cycles, etc. Of course, an above satisfactory work record was also a factor. However, prior to participation in the assessment center used to gather the dependent measure, the management backgrounds of all subjects were reviewed. This review examined:

organizational levels at which the subjects had management experience; number of functional areas in which the subjects had management experience; annual performance reviews of subjects; performance reviews related to subjects' participation on special projects or committees; educational level; training courses attended; special awards or commendations received by subjects; and subjects' participation in outside professional or civic organizations. Based on this review of the subjects' backgrounds, there was no indication that differences existed regarding the management experience or ability of experimental subjects versus control subjects.

Experimental Subjects

Experimental subjects were twelve managers employed by the organization who were trained and served as assessors in assessment programs conducted by the agency. All subjects were Grade Scale (GS) fourteen or fifteen and were applicants for the executive development program conducted by the agency. An overview of the executive development program follows in later portions of this section.

Control Subjects

Control subjects were sixty managers employed by the same organization who received no assessor training nor

served as assesors. All control subjects were a GS fourteen or fifteen and were applicants for the executive development program conducted by the agency.

Treatment

The format used in the assessor training session is outlined in Appendix B. This training, as well as the opportunity to serve as assessors in actual assessment centers, provided experimental subjects with an intimate understanding of the assessment center process. Specifically, each experimental subject was familiarized with the manner in which behavior exhibited in various simulation instruments was evaluated according to management skill categories, as well as, the basic concepts and rationale underlying the assessment center process.

The generic types of simulation instruments, description of the instruments, and management skills used in the assessor training appear in Appendix C, D, and E, respectively. As can be seen from Appendix C, one of the twelve experimental subjects did not receive assessor training on all of the simulation instruments comprising the treatment condition. This subject received training on only the in-basket and competitive group discussion instruments; however, the assessor training received by the subject

encompassed all of the management skills used in the experimental training session.

Dependent Measurement

The dependent measurement used in this study was performance in the executive development program assessment center. This assessment center was comprised of five simulation instruments, each of which is described in Appendix F. As can be seen from a comparison of the simulation instruments used in the training of experimental subjects with the instruments used in gathering the dependent measurement, a substantial degree of similarity existed in terms of the types of simulation instruments used; the exception being the inclusion of a cooperative group discussion in the executive development program assessment center. Experimental subjects received no assessor training on this type of simulation instrument; however, the task demands of the cooperative versus competitive group discussion were very similar. Both exercises were relatively unstructured situations in which group members were free to structure the discussion as they deemed

*Note: for the sake of clarity, skill names listed in Appendix E have been standardized. In some instances skill names were not consistent between the dependent measurement and treatment condition. For example, Interpersonal was sometimes labeled Sensitivity; however, definitions of the skills were the same and skill labels were interchangeable.

appropriate. The primary difference between the two exercises was in the degree of competition built into the exercises. In the competitive group discussion, each group member was assigned a particular position which they were instructed to advocate to the group at large. In the cooperative group discussion, group members were not assigned a particular position to advocate or defend.

The management skills on which subjects were evaluated are listed in Appendix G. A comparison of these skills with the management skills utilized in the assessor training session for experimental subjects demonstrates substantial overlap. With the exception of the skill of Adaptability, skills on which experimental subjects were trained to assess were identical to those skills which were assessed for the dependent measurement. Experimental subjects were not familiarized with this skill during the assessor training session.

The dependent measurement, performance in the executive development program assessment center, utilized well established assessment center procedures. Subjects completed the simulation instruments during a two day period. Another day was devoted to the assessor consensus meeting to arrive at an overall evaluation of each skill category. In addition to an overall evaluation of each skill category, a single numerical score was assigned to each subject

to indicate their overall level of performance in the center. This overall performance score was a function of overall performance in the individual skill categories and was achieved through a clinical, yet structured, combining of these skill category scores by the assessment staff. Weightings as to the importance of the skills measured by the assessment center for success on the actual job were provided to the assessment staff. These weightings were obtained by responses to a questionnaire which was distributed to incumbents of target positions within the organization. To determine the overall performance scores of assessees, all assessors comprising the assessment staff reviewed the performance of each assessee on the individual skill categories and then, utilizing the aforementioned weightings, arrived at a group consensus regarding the overall performance score of each assessee. The evaluation of subjects was based on a seven point numerical scale which appears in Appendix H. Because of the manner in which the overall assessment center performance evaluations were derived, the skill of adaptability was included in the analysis despite the fact that no performance differences were predicted for the skill since the experimental treatment did not include this skill category.

The assessment staff which observed and evaluated subjects was composed of individuals well versed in assessment

principles. The staff consisted of professional consultants in the area of assessment center technology, non-government managers with assessor experience and internal managers.

Statistical Analysis

To investigate the performance of the control and experimental group's on overall assessment center performance evaluations (Hypothesis 1), a one-way analysis of variance was applied to these evaluations.

To investigate the performance of the control and experimental group's on overall skill cagetory evaluations for each skill category (Hypothesis 2), nine separate 2 x 6 analyses of variance with repeated measures on the second factor were performed. In this design the first factor was the levels of treatment condition; experimental and control group. The second factor was simulation instrument type and overall skill category evaluation; in-basket, problem solving, cooperative group discussion, competitive group discussion, leadership simulation instrument and overall performance evaluations for the skill category. Such an analysis was performed for each of the nine skill categories assessed in the executive development program assessment center.

To investigate the performance of the control and experimental group's relative to overall performance on the

individual simulation instruments used in the assessment center (Hypothesis 3), a 2 x 5 analysis of variance with repeated measures on the second factor was applied to the data. The first factor was levels of the treatment condition; experimental and control group. The second factor was overall performance evaluations on each simulation instrument; in-basket, problem solving, cooperative group discussion, competitive group discussion, and leadership simulation instrument. Since subjects were not assigned an overall evaluation summarizing their performance for a given simulation instrument, scores for each subject were derived in the following manner. Ratings of the assessors for each skill category within a given simulation instrument were averaged. These mean values then became the rating values for the individual skill categories measured. The figures for each skill category were summed and divided by the number of skill categories measured by the instrument. The resulting quotient then served as the subjects' overall performance evaluation for a given simulation instrument. Thus, all skills measured within a given simulation instrument were weighted equally in arriving at the overall performance evaluation for the instruments.

Results

Tables of means, as well as, analysis of variance tables resulting from the statistical analyses are presented in Appendix I. Although the analysis of variance tables related to the nine individual skill categories presents results for the factor of Simulation Instrument Type and Overall Skill Category, the purpose of the study was to investigate performance of the experimental group relative to that of the control group. Thus, only findings related to the factors of Experimental Versus Control subjects and interaction effects are discussed relative to hjypothesis number two. Likewise, for hypothesis number three only findings relevant to the factors of Experimental versus Control subjects and interaction effects are discussed for similar reasons.

Hypothesis number one predicted no differences between the control and experimental groups relative to overall assessment center performance evaluations. As can be seen from Table 4, this hypothesis was supported since no significant differences between the two groups emerged, $\underline{F}$ (1) = 2.22, $p > .05$.

Hypothesis number two predicted that no differences would emerge with respect to the performance of the control group versus the performance of the experimental group on

any of the nine overall skill category evaluations. As can be seen from Tables 5-13, this hypothesis was generally supported. While significant differences between groups or interaction effects failed to emerge for the skill categories of leadership, sensitivity, decisiveness, organizing and planning, perception, decision making, oral communication, and written communication, significant differences between groups did emerge on the skill category adaptability, $\underline{F}$ (1) = 7.08, $p<.01$. Thus, the experimental group performed significantly higher than did the control group. In summary, hypothesis number two was generally supported since eight of the nine tests for main effects were not significant and all nine of the tests for interaction failed to produce significant findings.

Hypothesis number three predicted that no performance differences would emerge with respect to the control and experimental groups' overall performance on any of the individual simulation instruments. As can be seen from Table 14, this hypothesis was supported as no significant differences emerged between groups, $\underline{F}$ (1) = 3.67, $p>.05$ nor was there a significant interaction effect, $\underline{F}$ (4) = 0.39 $p>.05$.

Hypothesis number four predicted that no interaction effects would emerge between treatment groups and type of

simulation instrument for any of the skill categories assessed. As can be seen from Tables 4-13, no significant interaction effects emerged relative to any of the skill categories and thus the hypothesis was supported.

Discussion

Generally speaking the results of the study agreed with the general hypotheses of the experimentor. Hypothesis 1 was supported as no significant differences between groups emerged with respect to overall assessment center performance evaluations. Hypothesis 2 was also generally supported. Experimental subjects performed better than control subjects on only the skill category of adaptability. On all other skill categories, no differences between groups nor interaction effects emerged. Hypothesis 3 was supported as no differences between groups emerged with respect to overall performance on the individual simulation instruments, nor were there any significant interaction effects. Hypothesis 4 was supported as no significant interaction effects emerged between groups and type of simulation instrument for any of the skill categories assessed. In other words, the treatment condition failed to produce any effects on subsequent assessment center performance; with the possible exception being the performance of subjects on the skill category of adaptability. These findings are inconsistent with those of Bass (1954) and Petty (1974). Possible explanations are discussed below.

The study reported here differed significantly from these earlier studies with respect to the dependent measurement used. Both Petty and Bass used performance on a lone simulation instrument as the dependent measure. The study reported here used performance outcomes which were derived from subjects' performance across multiple simulation instruments. Thus one possible interpretation could be that any effects due to prior exposure with the assessment center process were minimized through the use of multiple instruments. In other words, "coaching" effects were difficult for subjects to maintain across the series of simulation instruments. However, it should be noted that no differences emerged between groups with regard to performance across the different types of simulation instruments. This finding appears to directly conflict with the finding of Bass and Petty, who both utilized a leaderless group discussion exercise to obtain dependent measures.

Another difference between the study reported here and the work of Bass and Petty lay in the makeup of the observers or assessors who evaluated subjects in the simulation instruments. In this study, assessors were well trained and experienced in evaluating the performance of assessment center participants. Earlier studies did not rely on well trained or experienced assessors to evaluate subjects. Thus another possible interpretation of the study's

differences in findings may be that well trained and experienced assessors were more capable of differentiating "faking" attempts of assessment center participants from behaviors which truly indicated effective managerial skills.

Interpretation of the fact that adaptability emerged as significantly different between groups is difficult since experimental subjects were not familiarized with the skill category during training. One possible interpretation could rest in the fact that experimental subjects actually served as assessors. While serving as assessors, subjects participated in team meetings with other assessors to discuss assessee performance and achieve a concensus regarding evaluations. Such meetings strongly emphasize cooperation and flexibility on the part of all assessors involved. Thus it could be argued that participation in such meetings led to an increase in the adaptability skills of the experimental subjects.

Another explanation could be that experimental subjects differed from control subjects on this skill prior to attending assessor training. As was pointed out earlier, appointment of subjects to attend assessor training was partially based on receptiveness to the assessment center process. Given this it could be argued that selected subjects were more receptive to new ideas and methods than

were other subjects, and that these characteristics are positively associated with the skill of adaptability. However, both of these interpretations must be viewed very tentatively. The fact that such an outcome was observed may well have been due to chance.

The results of this study also directly bear on claims regarding the "fringe" benefits derived by individuals who experience assessor training. While it seems likely that such individuals could benefit from improvement in interviewing or performance appraisal skills, improvement in management skills as measured by the assessment center process did not emerge as one of these "fringe" benefits; the possible exception to this being improvement in the skill category of adaptability. Thus while it could be argued that experiencing assessor training may sensitize trainees to their developmental needs, this alone does not lead to improvement nor guarantee that the individuals will seek out training to correct deficiencies.

Given the objectives of assessor training, to train individuals to observe, record, categorize, and evaluate assessee's behavior; the fact that such training did not improve the management skills measured in the assessment center process appears reasonable. Simply stated, training in the skills required to evaluate the managerial abilities of others should not be viewed as corresponding to

improvement of the managerial abilities of those who receive such training. More specifically, as assessor training does not emphasize direct practice and feedback in the various skill categories measured by the assessment center process, such training results in no improvement in these skills for individuals who undergo such training. Again, the possible exception being improvement in the skill category of adaptability.

In addition to these interpretations some consideration must also be given to the fact that since the study reported here took place in an actual operational assessment center, certain variables could not be controlled.

Experimental subjects did not experience identical assessor training sessions. Although the training format for assessor training sessions was uniform, subjects were trained to serve as assessors for various and different assessment centers conducted by the agency. Some subjects received assessor training for assessment centers targeted at first line management positions, while others received training for assessment centers targeted at mid-level management positions. Therefore the particular simulation instruments on which subjects received assessor training varied although the generic type of simulation instruments on which subjects were trained was highly uniform across sessions as was discussed earlier. Thus, with respect to

the organizational level of the instruments, there was not a large degree of similarity between the instruments used in the assessor training sessions and the instruments used in the gathering of the dependent measurement. Because of this, the ability of the experimental subjects to generalize from the assessor training sessions to the dependent measurement situation may have been significantly decreased.

The second variable which could not be controlled was the length of time which elapsed between the assessor training sessions for experimental subjects and the time of the dependent measurement. Eleven of the twelve experimental subjects experienced the treatment condition and actually served as an assessor a maximum of twelve months before being assessed in the executive development program assessment center. For one subject, the time between experiencing the treatment condition and being assessed in the executive development program assessment center was somewhat longer. In the studies of Bass and Petty, the dependent measurement was obtained more immediately after the coaching sessions. Thus, one interpretation of the study reported here could be that any improvement on assessment center performance attributable to prior exposure or coaching influence is short lived in nature and dissipates quickly with the passage of time.

Although the results of any one study must be viewed tentatively, the findings of this study indicate that prior exposure to the assessment center process, in the form of assessor training, has little or no impact on subsequent assessment center performance. As was indicated earlier, there is little research on prior exposure to the assessment center and its effects on subsequent assessment center performance. Certainly this appears to be an area that requires more investigation. Such studies could include replication of the study presented here under more controlled conditions utilizing pre- and post-test measures of subjects' ability levels in the various management skill categories. Also, investigations concerning the time interval, between assessor training and subsequent assessment of subjects seems called for, as are investigations to explore the relationship between the organizational level for which subjects are trained to serve as assessors and the organizational level at which they are subsequently assessed.

APPENDIX A

Skills Categories Measured Within Individual Simulation Instruments Comprising the Executive Development Program Assessment Center

Skill Category	Simulation Instruments				
	In-Basket	Problem Solving	Cooperative Leaderless Group Discussion	Competitive Leaderless Group Discussion	Leadership
Leadership	X	X	X	X	X
Sensitivity	X	X	X	X	X
Decision Making	X	X	X	X	X
Decisiveness	X	X	X	X	X
Organizing and Planning	X	X	X	X	X
Written Communication	X	X	O	O	O
Oral Communication	X	X	X	X	X
Adaptability	O	X	X	X	X
Perception	X	X	X	X	X

X = Skill Category Measured by Simulation Instrument
O = Skill Category Not Measured by Simulation Instrument

APPENDIX B

Training Schedule Used for Experimental Treatment

ASSESSOR TRAINING SCHEDULE

Day One

9:00 - 11:00	Background to Assessment Centers - History and Background
11:00 - 11:30	Distinguishing Conclusions from Support Statements
11:30 - 12:30	LUNCH
12:30 - 1:30	Review Skills and Weighting Sheet, Take Skill Categorization Exercise
1:30 - 2:45	Take Leaderless Group Discussion
2:45 - 3:45	Review Assessor Evaluation Guide for Leaderless Group Discussion
3:45 - 4:30	Practice Report Writing on Leaderless Group Discussion

Homework

1. Review Distinguishing Conclusion versus Support Statements Exercise
2. Review Skill Categorization Exercise

Day Two

9:00 - 9:30	Questions on Homework and Leaderless Group Discussion
9:30 - 10:45	Take In-Basket
10:45 - 11:15	Discuss In-Basket Interview

Day Two Continued

11:15 - 12:15	Review In-Basket Exercise and Guide
12:15 - 1:00	LUNCH
1:00 - 1:45	Do In-Basket Interview
1:45 - 2:00	Discuss Leadership Exercise
2:00 - 2:50	Take Leadership Exercise
2:50 - 3:45	Discuss Assessor Guide for Leadership and Role-Playing Instructions
3:45 - 4:30	Practice Report Writing on Leadership and In-Basket Exercises

Homework

1. Complete practice Reports on Leadership and In-Basket Exercises

Day Three

9:00 - 10:00	Review Problem Solving Exercise
10:00 - 10:30	Problem Solving Interview
10:30 - 11:00	Discuss Problem Solving
11:00 - 12:00	Discuss In-Basket and Leadership Exercises Practice Reports
12:00 - 1:00	LUNCH
Remainder of Day	Practice Interviewing and Rating of Skills

Day Three Continued

Homework

1. Review All Assessor Guides and Role-Playing Instructions

Day Four

Observe Practice Assessees

Homework

1. Complete Exercise Reports on Practice Assessees

Day Five

9:00 - 10:00	General Feedback on Practice Reports and Questions
10:00 - 12:00	Team Meeting
12:00 - 1:00	Discuss Final Reports

APPENDIX C

Types of Simulation Instruments on Which Experimental Subjects Were Trained as Assessors

Subject	Exercise Type				
	In-Basket	Problem Solving	Cooperative Group Discussion	Competitive Group Discussion	Leadership
1	X	X	O	X	X
2	X	X	O	X	X
3	X	X	O	X	X
4	X	X	O	X	X
5	X	O	O	X	O
6	X	X	O	X	X
7	X	X	O	X	X
8	X	X	O	X	X
9	X	X	O	X	X
10	X	X	O	X	X
11	X	X	O	X	X
12	X	X	O	X	X

X = Trained on Exercise Type
O = Untrained on Exercise Type

APPENDIX D

Overview of Simulation Instrument Types Used in Experimental Training Session

In-Basket Exercise

The assessee is required to assume the role of a hypothetical person in an organization and is given a series of memos, letters, project reports, telephone messages, etc. that would typically be found in this person's in-basket. Each assessee is given a specific amount of time to review the materials and to take appropriate action on them by writing letters, memos and notes to self, subordinates or superiors. After completing the in-basket, the assessee is interviewed regarding the approach to the task, the rationale for taking the actions indicated, and the opinions developed regarding subordinates, peers, supervisors and the organization.

Problem Solving Exercise

Each assessee is individually required to process a considerable amount of data regarding several alternatives regarding policy matters or other issues and is required to assimilate the data and arrive at a written recommendation and justification as to which of the several alternatives is superior. Based on the above described written report, the assessee is required to orally summarize and defend the recommendation to two assessors who assume specific roles; an acquiescent supportive person and a probing, challenging individual.

Competitive Group Discussion Exercise

Six to Seven assessees are placed in an unstructured group situation with general background data provided to each of them. Each assessee is also given data outlining the merits of his/her particular assigned position. Each member is then given an opportunity to orally describe the merits of his/her position after which the group engages in a discussion ultimately arriving at a group recommendation about which of the positions is most meritorious.

Leadership Exercise

The assessee is provided with a packet of materials and instructions regarding a task which must be performed within a specified period of time. After the assessee has read the instructions regarding the nature of the task that is required, two assistants are introduced into the exercise situation. The assessee is free to make whatever use of the two assistants in the accomplishment of the task as he/she deems appropriate. The two assistants are role players; one of whom assumes an agreeable but somewhat inept role, while the other assumes a disagreeable attitude but is able to perform work effectively. Although the assessee is given formal authority over the two assistants,

whenever possible the disagreeable but competent assistant attempts to assume the leadership role in the work group.

APPENDIX E

Managerial Skills Used in Assessor Training Experience Attended by Experimental Subjects

Organizing and Planning: Effective in arranging and relating work. Effectively plans and organizes own activities as well as those of a group. Establishes well-defined work objectives and priorities for accomplishing them.

Perception: Identifies the factors essential to the solution of a problem. The ability to seek out pertinent data and put it together in order to solve a problem either in a group or individually.

Decision Making: The ability to reach logical conclusions and make decisions on evidence at hand.

Decisiveness: The ability to make decisions when required and to take action when appropriate.

Leadership: Give direction to and coordinate the activities of others. The ability to lead a group to accomplish a task and get ideas accepted.

Sensitivity: The ability to deal effectively with people and have one's ideas acted upon in a positive manner.

Oral Communication: The ability to communicate effectively through the use of oral skills.

Written Communication: The ability to communicate effectively through writing.

APPENDIX F

Overview of Simulation Instrument Types Used in Executive Development Program Assessment Center

In-Basket Exercise

The assessee is required to assume the role of a hypothetical person in an organization and is given a series of memos, letters, project reports, telephone messages, etc. that would typically be found in this person's in-basket. Each assessee is given a specific amount of time to review the materials and to take appropriate action on them by writing letters, memos and notes to self, subordinates or superiors. After completing the in-basket, the assessee is interviewed regarding the approach to the task, the rationale for taking the actions indicated, and the opinions developed regarding subordinates, peers, supervisors and the organization.

Problem Solving Exercise

Each assessee is individually required to process a considerable amount of data regarding several alternatives regarding policy matters or other issues and is required to assimilate the data and arrive at a written recommendation and justification as to which of the several alternatives is superior. Based on the above described written report, the assessee is required to orally summarize and defend the recommendation to two assessors who assume specific roles; an acquiescent supportive person and a probing, challenging individual.

Competitive Group Discussion Exercise

Six to Seven assessees are placed in an unstructured group situation with general background data provided to each of them. Each assessee is also given data outlining the merits of his/her particular assigned position. Each member is then given an opportunity to orally describe the merits of his/her position after which the group engages in a discussion ultimately arriving at a group recommendation about which of the positions is most meritorious.

Leadership Exercise

The assessee is provided with a packet of materials and instructions regarding a task which must be performed within a specified period of time. After the assessee has read the instructions regarding the nature of the task that is required, two assistants are introduced into the exercise situation. The assessee is free to make whatever use of the two assistants in the accomplishment of the task as he/she deems appropriate. The two assistants are role players; one of whom assumes an agreeable but somewhat inept role, while the other assumes a disagreeable attitude but is able to perform work effectively. Although the assessee is given formal authority over the two assistants,

whenever possible the disagreeable but competent assistant attempts to assume the leadership role in the work group.

Cooperative Group Discussion Exercise

Six to seven assessees are place in an unstructured group situation with general background data provided to each of them. The task of the group is to achieve a consensus of opinion regarding a solution to the problem which is posed in the exercise. Unlike the competitive group discussion, assessees are not assigned a particular position to advocate or defend, but rather seek a solution to the problem in a cooperative setting.

APPENDIX G

Managerial Skills Assessed in the Executive Development Program Assessment Center

Organizing and Planning: Effective in arranging and relating work. Effectively plans and organizes own activities as well as those of a group. Establishes well-defined work objectives and priorities for accomplishing them.

Perception: Identifies the factors essential to the solution of a problem. The ability to seek out pertinent data and put it together in order to solve a problem either in a group or individually.

Decision Making: The ability to reach logical conclusions and make decisions on evidence at hand.

Decisiveness: The ability to make decisions when required and to take action when appropriate.

Leadership: Give direction to and coordinate the activities of others. The ability to lead a group to accomplish a task and get ideas accepted.

Sensitivity: The ability to deal effectively with people and have one's ideas acted upon in a positive manner.

Oral Communication: The ability to communicate effectively through the use of oral skills.

Written Communication: The ability to communicate effectively through writing.

Adaptability: Effectively modifying one's behavior as a function of situational changes as well as persons with whom one interacts.

APPENDIX H

Numerical Evaluation Scale Used in the Executive Development Program Assessment Center

7 - Outstanding

6 - Well Above Satisfactory

5 - Above Satisfactory

4 - Satisfactory

3 - Below Satisfactory

2 - Well Below Satisfactory

1 - Weak

APPENDIX I

Tables of Means and Analysis of Variance Results

TABLE 1

Means of Experimental and Control Subjects on Overall Assessment Center Performance Evaluations

	MEAN
Experimental Subjects	4.50
Control Subjects	4.15

TABLE 2

Mean Evaluations of Experimental and Control Subjects for Each Skill Category Across Simulation Instrument Types and Overall Skill Category Evaluations

	In-Basket	Problem Solving	Cooperative Leaderless Group Discussion	Competitive Leaderless Group Discussion	Leadership	Overall Skill Category Evaluation
Leadership						
Experimental Subjects	5.25	0	4.42	4.08	3.75	4.08
Control Subjects	4.90	0	4.05	3.80	3.52	3.98
Sensitivity						
Experimental Subjects	5.00	4.50	4.83	4.17	4.67	4.58
Control Subjects	4.73	4.55	4.25	4.13	4.23	4.33
Decision Making						
Experimental Subjects	5.17	4.75	4.25	4.17	3.75	4.50
Control Subjects	4.48	4.33	4.35	4.18	3.63	4.15
Decisiveness						
Experimental Subjects	5.50	5.42	4.92	4.42	4.25	5.00
Control Subjects	4.73	4.70	4.35	4.42	3.92	4.40
Organizing and Planning						
Experimental Subjects	5.08	5.08	4.67	4.00	3.50	4.50
Control Subjects	5.05	4.37	4.37	4.23	3.55	4.33
Written Communication						
Experimental Subjects	4.67	4.50	0	0	0	4.58
Control Subjects	4.42	4.00	0	0	0	4.27
Oral Communication						
Experimental Subjects	4.08	4.67	4.67	4.33	4.25	4.58
Control Subjects	4.48	4.32	4.40	4.25	4.08	4.30
Adaptability						
Experimental Subjects	0	5.00	4.42	4.42	3.92	4.50
Control Subjects	0	3.77	4.03	3.87	3.50	3.87
Perception						
Experimental Subjects	4.50	5.00	4.83	4.58	3.92	4.67
Control Subjects	4.23	4.48	4.52	4.43	3.63	4.33

TABLE 3

Mean Evaluations of Experimental and Control Subjects for Overall Performance on Each Simulation Instrument

	Experimental Subjects	Control Subjects
In-Basket	4.96	4.65
Problem Solving	4.87	4.33
Cooperative Leaderless Group Discussion	4.62	4.26
Competitive Leaderless Group Discussion	4.30	4.15
Leadership	4.01	3.78

TABLE 4

Analysis of Variance Results Comparing Experimental and Control Group on Overall Assessment Center Performance

Source of Variation	df	MS	F
(A) Experimental vs. Control	1	1.23	2.22
(B) $Error_w$	70	.55	

* $p<.05$

TABLE 5

Analysis of Variance Results Comparing Performance of Experimental and Control Group on the Skill Category of Leadership

Source of Variation	df	MS	F
(A) Experimental vs. Control	1	2.96	0.82
(B) Simulation Instrument Type and Overall Skill Category[1]	5	217.11	218.97**
AXB	5	0.21	0.21
$Error_b$	70	3.60	
$Error_w$	350	0.99	

* $p<.05$
** $p<.01$

1 In-basket vs. Problem Solving vs. Cooperative Leaderless Group Discussion vs. Competitive Leaderless Group Discussion vs. Leadership vs. Overall Skill Category

TABLE 6

Analysis of Variance Results Comparing Performance of Experimental and Control Group on the Skill Category of Sensitivity

Source of Variation		df	MS	F
(A)	Experimental vs. Control	1	3.83	1.66
(B)	Simulation Instrument Type and Overall Skill Category[1]	5	3.49	4.54**
	AXB	5	0.56	0.73
	$Error_b$	70	2.31	
	$Error_w$	350	0.77	

* $p<.05$
** $p<.01$

[1] In-basket vs. Problem Solving vs. Cooperative Leaderless Group Discussion vs. Competitive Leaderless Group Discussion vs. Leadership vs. Overall Skill Category

TABLE 7

Analysis of Variance Results Comparing Performance of Experimental and Control Group on the Skill Category of Decisiveness

Source of Variation		df	MS	F
(A)	Experimental vs. Control	1	14.83	3.38
(B)	Simulation Instrument Type and Overall Skill Category[1]	5	7.50	6.66**
	AXB	5	0.82	0.73
	$Error_b$	70	4.39	
	$Error_w$	350	1.13	

* $p<.05$

** $p<.01$

[1] In-basket vs. Problem Solving vs. Cooperative Leaderless Group Discussion vs. Competitive Leaderless Group Discussion vs. Leadership vs. Overall Skill Category

TABLE 8

Analysis of Variance Results Comparing Performance of Experimental and Control Group on the Skill Category of Organizing and Planning

Source of Variation		df	MS	F
(A)	Experimental vs. Control	1	1.45	0.04
(B)	Simulation Instrument Type and Overall Skill Category[1]	5	17.25	20.20**
	AXB	5	1.09	1.27
	$Error_b$	70	3.07	
	$Error_w$	350	0.85	

* $p<.05$
** $p<.01$

1 In-basket vs. Problem Solving vs. Cooperative Leaderless Group Discussion vs. Competitive Leaderless Group Discussion vs. Leadership vs. Overall Skill Category

TABLE 9

Analysis of Variance Results Comparing Performance of Experimental and Control Group on the Skill Category of Adaptability

Source of Variation	df	MS	F
(A) Experimental vs. Control	1	15.00	7.08**
(B) Simulation Instrument Type and Overall Skill Category[1]	5	169.57	168.99**
AXB	5	2.18	2.18
$Error_b$	70	2.12	
$Error_w$	350	1.00	

* $p<.05$

** $p<.01$

1 In-basket vs. Problem Solving vs. Cooperative Leaderless Group Discussion vs. Competitive Leaderless Group Discussion vs. Leadership vs. Overall Skill Category

TABLE 10

Analysis of Variance Results Comparing Performance of Experimental and Control Group on the Skill Category of Perception

Source of Variation	df	MS	F
(A) Experimental vs. Control	1	5.81	2.49
(B) Simulation Instrument Type and Overall Skill Category[1]	5	8.05	7.72**
AXB	5	0.14	0.14
$Error_b$	70	2.33	
$Error_w$	350	1.04	

* $p<.05$
** $p<.01$

1 In-basket vs. Problem Solving vs. Cooperative Leaderless Group Discussion vs. Competitive Leaderless Group Discussion vs. Leadership vs. Overall Skill Category

TABLE 11

Analysis of Variance Results Comparing Performance of Experimental and Control Group on the Skill Category of Decision Making

Source of Variation	df	MS	F
(A) Experimental vs. Control	1	3.50	1.43
(B) Simulation Instrument Type and Overall Skill Category[1]	5	7.37	7.87**
AXB	5	0.87	.93
$Error_b$	70	2.46	
$Error_w$	350	0.94	

* $p<.05$
** $p<.01$

1 In-basket vs. Problem Solving vs. Cooperative Leaderless Group Discussion vs. Competitive Leaderless Group Discussion vs. Leadership vs. Overall Skill Category

TABLE 12

Analysis of Variance Results Comparing Performance of Experimental and Control Group on the Skill Category of Oral Communication

Source of Variation		df	MS	F
(A)	Experimental vs. Control	1	0.94	0.62
(B)	Simulation Instrument Type and Overall Skill Category[1]	5	1.08	2.52*
	AXB	5	0.75	1.75
	$Error_b$	70	1.52	
	$Error_w$	350	0.43	

* $p<.05$
** $p<.01$

[1] In-basket vs. Problem Solving vs. Cooperative Leaderless Group Discussion vs. Competitive Leaderless Group Discussion vs. Leadership vs. Overall Skill Category

TABLE 13

Analysis of Variance Results Comparing Performance of Experimental and Control Group on the Skill Category of Written Communication

Source of Variation		df	MS	F
(A)	Experimental vs. Control	1	1.61	1.93
(B)	Simulation Instrument Type and Overall Skill Category[1]	5	393.79	991.65**
	AXB	5	0.52	1.30
	$Error_b$	70	0.84	
	$Error_w$	350	0.40	

* $p<.05$
** $p<.01$

1 In-basket vs. Problem Solving vs. Cooperative Leaderless Group Discussion vs. Competitive Leaderless Group Discussion vs. Leadership vs. Overall Skill Category

TABLE 14

Analysis of Variance Results Comparing Overall Performance of Experimental and Control Group on Simulation Instruments

Source of Variation		df	MS	F
(A)	Experimental vs. Control	1	5.105	3.67
(B)	Simulation Instrument Type[1]	4	7.56	14.41**
	AXB	4	0.21	0.39
	$Error_b$	70	1.39	
	$Error_w$	280	0.52	

* $p<.05$
** $p<.01$

[1] In-basket vs. Problem Solving vs. Cooperative Leaderless Group Discussion vs. Competitive Leaderless Group Discussion vs. Leadership

References

Bass, B.M. The leaderless group discussion. Psychological Bulletin, 1954, 51, 465-492.

Bender, J.M. What is typical of assessment centers? Personnel, 1973, 50(4), 50-57.

Bray, D.W., Campbell, R.J., & Grant, D.L. Formative years in business: A long-term AT&T study of managerial lives. New York: Wiley, 1974.

Burroughs, W.A, Rollins, J.B., & Hopkins, J.J. The effects of age, departmental experience and prior rater experience on performance in assessment center exercises. Academy of Management Journal, 1973, 16, 335-339.

Byham, W.C. Assessment centers for spotting future managers. Harvard Business Review, 1970, 48 (4), 150-160.

Cohen, S.L. Standardization of assessment center technology: some critical concerns. Journal of Assessment Center Technology, 1978, 1(3), 1-10.

Denning, D.L., & Grant, D.L. Knowledge of the assessment process: does it influence candidate ratings? Journal of Assessment Center Technology, 1979, 2(2), 7-12.

Huck, J.R. Assessment centers: A review of the external and internal validities. Personnel Psychology, 1973, 26, 191-212.

Jaffee, C.L., & Michaels, C.E., Jr. Is in-basket performance subject to coaching effects? Journal of Assessment Center Technology, 1978, 1(1), 13-17.

Moses, J.L. The development of an assessment center for the early identification of supervisory potential. Personnel Psychology, 1973, 26, 569-580.

Petty, M.M. A multivariate analysis of the effects of experience and training upon performance in a leaderless group discussion. Personnel Psychology, 1974, 27, 271-282.